DOCUMENT RESUME

ED 306 278 TM 013 129

AUTHOR Saka, Thomas; And Others
TITLE Longitudinal Student Achievement Monitoring in Hawaii Public Schools.
PUB DATE Mar 89
NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; Achievement Tests; Elementary School Students; Elementary Secondary Education; Grade 3; Grade 6; Grade 8; Grade 10; High School Students; *Longitudinal Studies; Low Achievement; Mathematics Achievement; *Public Schools; Reading Achievement; *Standardized Tests; *State Programs; Testing Programs
IDENTIFIERS *Hawaii; *Longitudinal Student Achievement Monitoring; Stanford Achievement Tests

ABSTRACT

In response to state mandates, the Hawaii Department of Education (HDE) began reporting on the current achievement status and progress of students scoring in the lowest three stanines of the Stanford Achievement Tests (SAT). The Longitudinal Student Achievement Monitoring (LSAM) process is described and evaluated; it was established to monitor low achievers and, eventually, to monitor achievement of all students in the state. The LSAM is designed to analyze students' achievement test scores over 2- and 4-year periods. The SAT is administered annually to 12,000 students at each of four grade levels--grades 3, 6, 8, and 10. This longitudinal analysis used surviving cohorts of students over 2- and 4-year periods from the fall of 1983 to the spring of 1988. The analysis included approximately 7,000 students per grade level. Reading and mathematics test scores on the SAT were used to represent educational achievement in the study. Results of the longitudinal analysis indicate continuous gains in achievement from elementary to high school grade levels. The largest mean gains occur in reading achievement. Programs and services offered by the HDE appear to be the most effective with low achievers for whom achievement mean stanine gains of up to one stanine in both reading and mathematics were found. Six data tables and six graphs are included. (TJH)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED306278

# Longitudinal Student Achievement Monitoring In Hawaii Public Schools

by

Thomas Saka
University of Hawaii

Rose Yamada
Hawaii Department of Education

Morris Lai
University of Hawaii

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

THOMAS SAKA

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the annual meeting of the American Educational Research Association, March, 1989. San Francisco, CA.

TM013129

## Longitudinal Student Achievement Monitoring In Hawaii Public Schools

Beginning in 1983, the Hawaii Department of Education was required to report to the Governor and the State Legislature on the current achievement status and progress of students scoring in the lowest three stanines of the Stanford Achievement Test (SAT) administered statewide. The Longitudinal Student Achievement Monitoring (LSAM) process was initiated in partial response to this requirement. The first application of the process was focused on the low achievers, defined as students scoring in the below average stanines (1-3) of the SAT. Subsequently, the monitoring has been expanded to include the achievement progress of all students. Longitudinal analysis is necessary to insure that all schools are striving to improve their instructional programs regardless of the academic characteristics of their students.

### Background

Since the late seventies, major national and local efforts have been directed toward the improvement of basic skills. In Hawaii, the goal of the Department's overall basic skills improvement effort has been:

> " to equip each student with the skills, knowledge, attitudes and values necessary to give an individual responsible direction to his/her life. The focus of instruction is to assist individuals in the development of those basic reading, writing, listening, speaking mathematical and thinking skills necessary for effective communication with others" (DOE, 1983; p. 2).

Improvement of basic skills and accounting for students' achievement via test scores have become further entrenched in the public's eye with the national calls for reform and academic excellence. These conditions magnify the importance of the Hawaii Department of Education's direction to mount and sustain effective schools that contribute to increasing academic achievement for every student.

The advantages of a statewide school system include the uniform administration of a minimum testing program. This advantage needed to be capitalized on to further assist key decision makers in assessing the effectiveness of schools and improvements in the use of existing resources. Other uses of the test results such as reporting current status of students, diagnosing students' weaknesses through item analyses and determining gross trends of performance at each of the grade levels tested continue to be important. However, the longitudinal achievement monitoring was viewed as essential for decision makers to determine the current academic status of students <u>in relation to</u> their prior achievement levels.

### Purpose of the Study

The purpose of this study is to describe the method by which the actions and effectiveness of Hawaii's public schools, to improve students' achievement levels, are evaluated and the progress of the students are monitored. While basically simple, the analysis of students' achievement test scores over two- and four-year periods provides a means to assess the impact of instructional

services across achievement and grade levels. The need for longitudinal analysis is present because intervention services generally seek substantive change in student's performance, and these changes many not be evident as short-term gains.

In the event that educational interventions may differentially affect students, the analysis compared growth rates among the three achievement groupings. The procedures further allow for assessing the effect of background differences such as gender and ethnicity on achievement gains.

The study also presents problems which may arise when changing versions of the basic test and the effects on the analysis. In Hawaii public schools, the seventh edition of the SAT was adopted during the 1986-87 school year. The sixth edition was used in prior years.

**Method**

The Stanford Achievement Test is administered yearly to all Hawaii public school students in the third, sixth, eighth and tenth grades. Thus standardized test data are available for individual students, on average, every two years. Since the state of Hawaii is a unified school district, data are available for approximately 12,000 students in each grade level tested each year.

The longitudinal analysis utilized surviving cohorts of students over two- and four-year periods from the fall of 1983 to the spring of 1988. The analysis which focused on students who had complete test data in successive testing years (e.g., grade 6 during the 1984-85 school year and grade 8 during the 1986-87 school year) contained approximately 7,000 students per grade level analysis. It is estimated that there is an average transiency rate of 2,000 students per grade level annually.

The students were matched across years through a ten-digit identification number which is assigned to each student by the Department of Education upon entering the public school system.

The total reading and total mathematics subtest scores of the SAT are used to represent educational achievement in this study. The use of these scores of the test provide global data on two areas of achievement that are closely scrutinized in the educational setting.

The Department of Education (DOE) annually produces two reports:

1) Student Achievement Report: Low Achievers

2) Student Achievement Report: Basic Skills.

The measures used in these reports are percent of students scoring in each stanine group, scaled score changes and mean stanines. Mean stanines, while not the most psychometrically acceptable measure of reporting data of this nature, are used in this study because there use allows for the easiest understanding among policy makers who often have no measurement background and for consistency with the DOE reports.

## DISCUSSION / FINDINGS

### Overall Analysis Over 2 and 4 Years

Four different cohorts of students were followed over a two year period. The cohort characteristics consisted of a) grade movement from 6 to 8 and from grade 8 to 10; and b) over two sets of years from 1984 to 1986 and from 1985 to 1987. The year designation represents the beginning of the school year where the 1987 designation represents the 1987-88 school year. As shown in Table 1 and Figure 1 there were significant positive increases in mean stanine performance on both the reading and math subtests for the movement of students from grade 6 to grade 8 in both time cohorts. The same trend can be seen in Figure 2 for the grade movement from 8 to 10; however, the magnitude of change was not as great.

Table 1. Mean Stanine Gains: Cohort of Students from Grade 6 to Grade 8.

| Cohort | SAT Math | | | | SAT Reading | | | |
|---|---|---|---|---|---|---|---|---|
| | N | Pre Mean | Post Mean | Gain | N | Pre Mean | Post Mean | Gain |
| 1984 - 1986 | 7,826 | 4.79 | 5.19 | 0.40 | 7,951 | 4.25 | 5.17 | 0.93 |
| 1985 - 1987 | 7,388 | 4.91 | 5.13 | 0.22 | 7,482 | 4.32 | 4.90 | 0.59 |

Table 2. Mean Stanine Gains: Cohort of Students from Grade 8 to Grade 10.

| Cohort | SAT Math | | | | SAT Reading | | | |
|---|---|---|---|---|---|---|---|---|
| | N | Pre Mean | Post Mean | Gain | N | Pre Mean | Post Mean | Gain |
| 1984 - 1986 | 7,408 | 5.02 | 5.16 | 0.14 | 7,064 | 3.97 | 4.95 | 0.98 |
| 1985 - 1987 | 5,894 | 5.05 | 5.38 | 0.33 | 5,585 | 4.02 | 4.67 | 0.65 |

In Figures 3 and 4 and Table 3, the performance of students over a four-year period and three test administrations is displayed. The first cohort containing students moving from grades 4, 6 and 8 shows continual increases in performance in both achievement areas over the three testing periods. The same is true for the math subtest of the cohort of students followed from grades 6, 8 and 10. However, the grade 8 1985 testing showed a drop in mean reading stanine for the group.

Table 3. Mean Stanines Gains: Cohort of Students from 1983 to 1987.

| Cohort Grades | SAT Math N | SAT Math 1983 Mean | SAT Math 1985 Mean | SAT Math 1987 Mean | SAT Reading N | SAT Reading 1983 Mean | SAT Reading 1985 Mean | SAT Reading 1987 Mean |
|---|---|---|---|---|---|---|---|---|
| 4 - 6 - 8 | 6,294 | 4.34 | 4.97 | 5.17 | 6,376 | 4.07 | 4.36 | 4.93 |
| 6 - 8 - 10 | 5,144 | 4.98 | 5.10 | 5.40 | 5,081 | 4.29 | 4.06 | 4.71 |

Educators in the state of Hawaii have been led to believe that there is a student performance drop at the eighth grade. This belief has been verified by one cohort of students in reading achievement while being refuted by a second cohort. Future analyses need to be conducted to gather data to identify a consistent trend.

**Analysis by Achievement Ability**

An analysis of student achievement gains by pretest achievement level shows somewhat unexpected results. The students were placed in achievement level groups based on their pretest stanine score. Three groups were utilized: below average (stanines 1-3), average (stanines 4-6) and above average (stanines 7-9). The analysis was conducted to determine if the education program interventions were differentially affecting students across achievement levels.

The results of the analyses can be found in Tables 4 & 5 and Figures 5 & 6. As expected, the below average group obtained the highest overall mean gain scores over a two-year period. A point of interest is that the above average group had statistically significant reading gains from grade 6 to grade 8 for both cohorts of students.

Table 4. Mean Stanine Gains: Movement from Grade 6 to Grade 8 by Achievement Level.

| Cohort | SAT Math | | | | SAT Reading | | | |
|---|---|---|---|---|---|---|---|---|
| | N | Pre Mean | Post Mean | Gain | N | Pre Mean | Post Mean | Gain |
| Grade 6 1984 - Grade 8 1986 | | | | | | | | |
| Below Avg. | 2220 | 2.17 | 3.21 | 1.05 | 2151 | 2.30 | 3.40 | 1.10 |
| Average | 3920 | 5.02 | 5.30 | 0.28 | 5119 | 4.65 | 5.55 | 0.90 |
| Above Avg. | 1686 | 7.71 | 7.54 | -0.17 | 681 | 7.39 | 7.97 | 0.58 |
| Grade 6 1985 - Grade 8 1987 | | | | | | | | |
| Below Avg. | 1935 | 2.17 | 3.02 | 0.85 | 1883 | 2.31 | 3.18 | 0.87 |
| Average | 3737 | 5.05 | 5.09 | 0.04 | 4921 | 4.66 | 5.18 | 0.52 |
| Above Avg. | 1716 | 7.70 | 7.59 | -0.11 | 678 | 7.35 | 7.63 | 0.28 |

Table 5. Mean Stanine Gain: Movement from Grade 8 to Grade 10 by Achievement Level.

| Cohort | SAT Math | | | | SAT Reading | | | |
|---|---|---|---|---|---|---|---|---|
| | N | Pre Mean | Post Mean | Gain | N | Pre Mean | Post Mean | Gain |
| Grade 8 1984 - Grade 10 1986 | | | | | | | | |
| Below Avg. | 1714 | 2.16 | 2.99 | 0.84 | 2954 | 2.62 | 3.53 | 0.91 |
| Average | 4028 | 5.06 | 5.10 | 0.04 | 3500 | 4.52 | 5.70 | 1.18 |
| Above Avg. | 1666 | 7.88 | 7.55 | -0.33 | 610 | 7.31 | 7.52 | 0.20 |
| Grade 8 1985 - Grade 10 1987 | | | | | | | | |
| Below Avg. | 1337 | 2.16 | 3.22 | 1.07 | 2249 | 2.63 | 3.34 | 0.71 |
| Average | 3186 | 5.04 | 5.23 | 0.19 | 2836 | 4.54 | 5.30 | 0.76 |
| Above Avg. | 1371 | 7.88 | 7.84 | -0.05 | 500 | 7.29 | 7.05 | -0.24 |

The above average group obtained negative gains in mathematics achievement across grade level movements and student cohorts. In the movement of students from grade 8 to grade 10, both cohorts show that the average stanine group had larger mean gains than the below average stanine group on the reading subtest. The expected trend, as shown in the other analyses is for the lowest pretest achievement group to have the largest gains and the highest pretest achievement group to have the lowest gains (sometimes negative).

### Longitudinal Comparisons with Different Versions of a Test

As can be expected when using commercially developed instruments to collect data over extended periods of time, the test instruments often change to reflect changes in criteria. During the 1986-87 school year the seventh edition of the Stanford Achievement Test was utilized by the state of Hawaii. Prior to this period the sixth edition of the SAT was used. While changes in test content, difficulty and metric were made to the seventh edition of the SAT, the test publishers assured users that comparisons between the different editions of the test could still be made if the SAT version 6 scores were converted to equivalent SAT version 7 scores.

Table 6 displays the mean gains scores of both the reading and mathematics subtests for the last four test administrations. The seventh edition of the SAT was first given in the 1986-87 school year and affects comparisons starting with the 1984 - 1986 analyses. Prior to that year the SAT version 7 was used.

Table 6. Comparison of Gains Over Four Test Administrations.

| | Grade 6 -> 8 | | Grade 8 -> 10 | |
|---|---|---|---|---|
| | Math | Reading | Math | Reading |
| 1982 - 1984 | -.10 | -.36 | -.16 | .07 |
| 1983 - 1985 | .13 | .18 | -.24 | -.08 |
| 1984 - 1986 | .40 | .93 | .14 | .98 |
| 1985 - 1987 | .22 | .59 | .33 | .65 |

In order for a longitudinal analysis to take place between the 1984 scores and the 1986 scores, the 1984 scores needed to be converted to the SAT version seven metric. The same conversion was needed in the 1985 - 1987 analysis. As shown in Table 6 the mean gains scores resulting from the analyses conducted with the conversion (1984-86 & 1985-87) are substantially larger than analyses conducted across years that did not require the conversion (1982-84 & 1983-85).

DOE administrators have asserted that no significant changes were made to educational offerings or curriculum in the 1985-86 school year that could account for the increase in mean gain scores. Discussion of the possible factors of the conversion which may have caused an artificial increase in the scores is beyond the scope of this paper. However, one has to question the validity of analyses using different versions of the test.

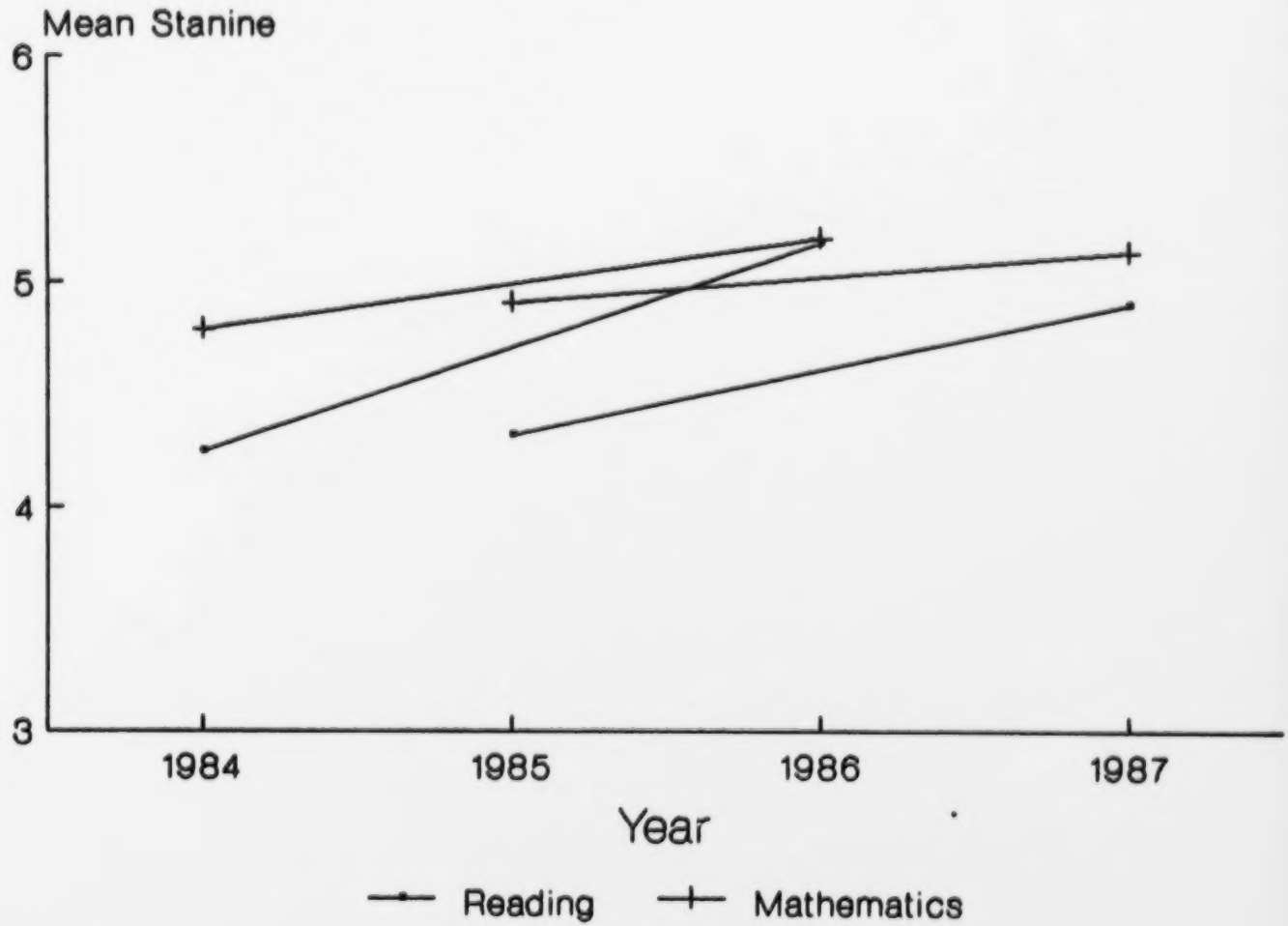
## Conclusions

Analysis of the achievement test scores of students who were administered the SAT for two- and four- year intervals shows continuous gains in achievement. This comes as a surprise in the state of Hawaii because of the popular belief that students "fall behind" in latter years of public schooling.

The study also shows that in general the progress of students appears to continually improve from elementary to high school grade levels. Although students start lower in reading as compared to mathematics, the largest mean gains occur in reading achievement. The programs or services offered by the department appear to be most effective with the low achievers (students scoring in stanines 1-3) where achievement mean stanine gains of up to 1 stanine in both reading and mathematics were found.

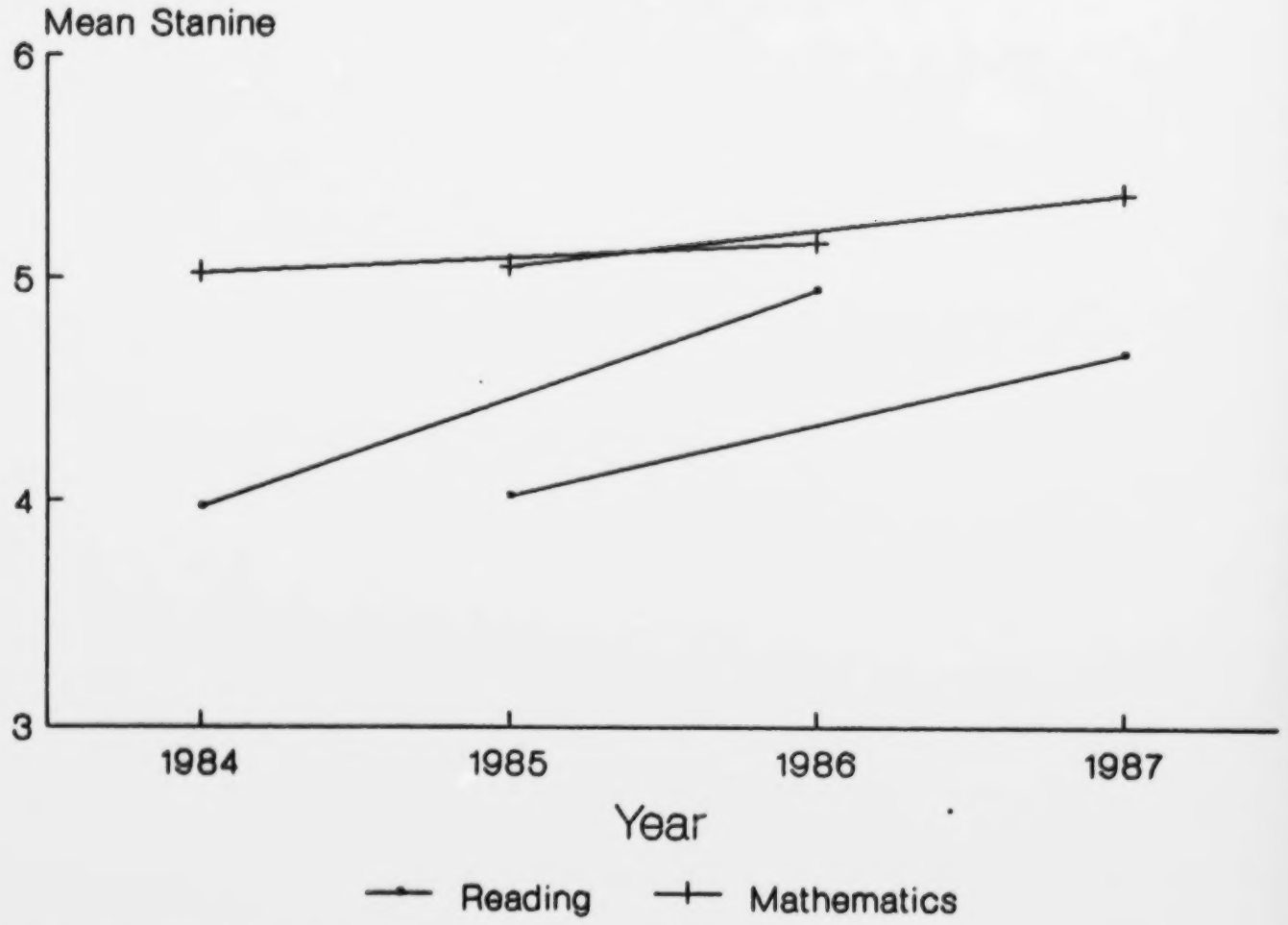
The high transiency rate of students in Hawaii public schools makes it difficult to determine the effects of the educational services for all students. However, sufficiently high proportions of students are studied to determine areas of programmatic strengths/weaknesses.

**Figure 1.**
**Surving Cohorts - Grades 6 to 8**



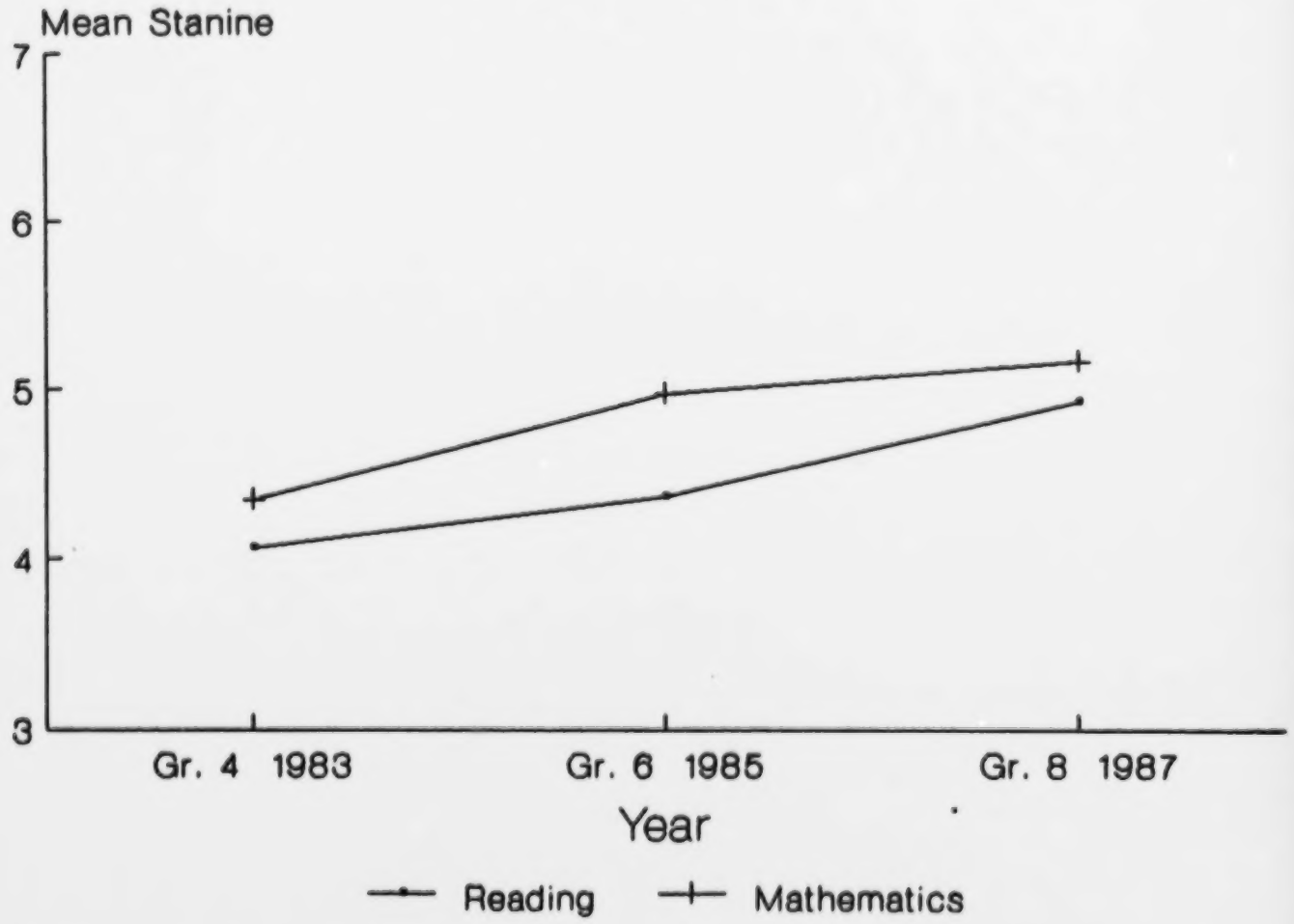

Stanford Achievement Test

**Figure 2.**
**Surving Cohorts - Grades 8 to 10**



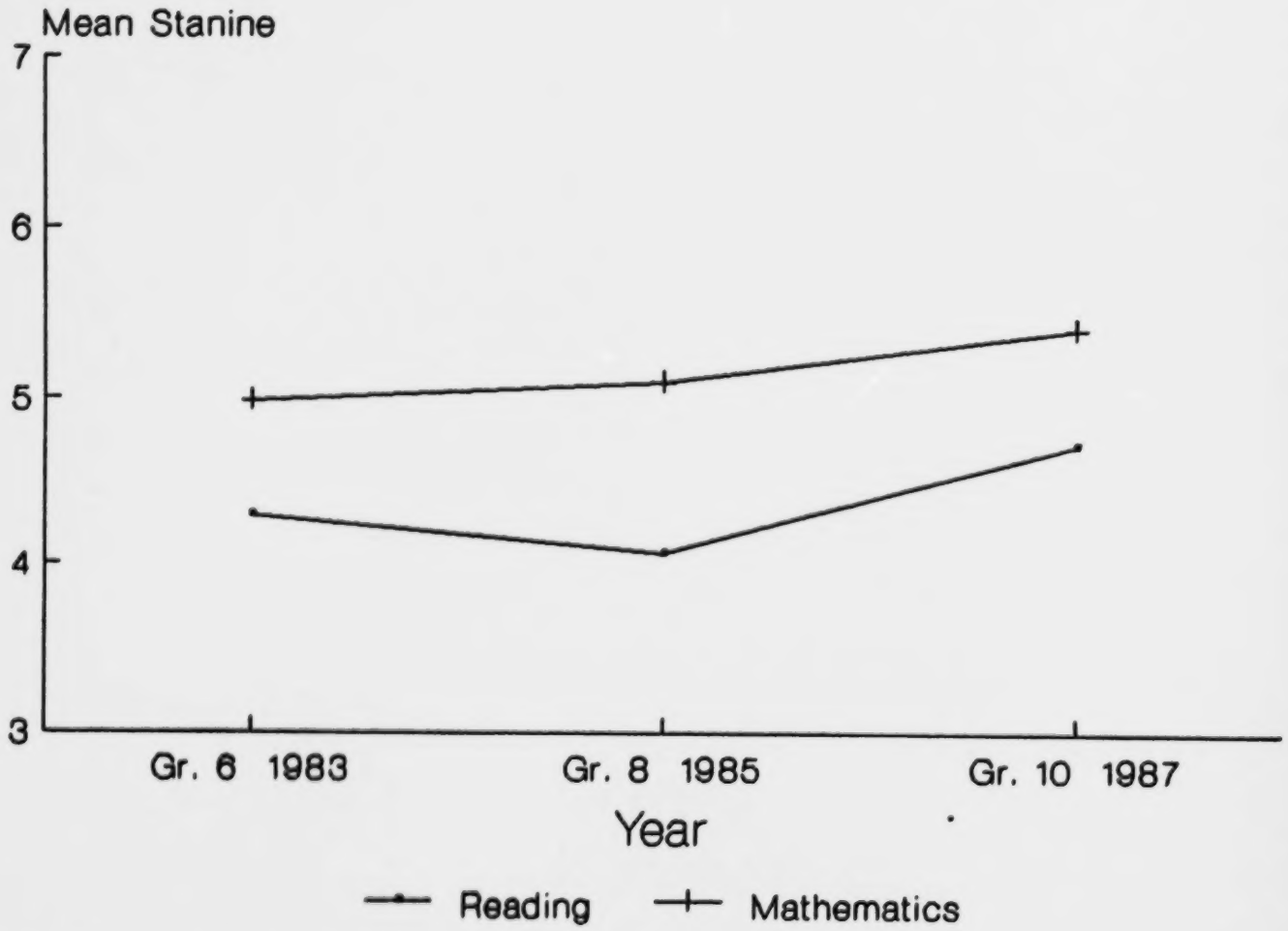

Stanford Achievement Test

**Figure 3.**
**Surving Cohorts Over Four Years**




Stanford Achievement Test N=6294

**Figure 4.**
**Surving Cohorts Over 4 Years**




**Stanford Achievement Test N=5144**

**Figure 5. Mean Gains by Pretest Achievement Level: Grade 6 to Grade 8**

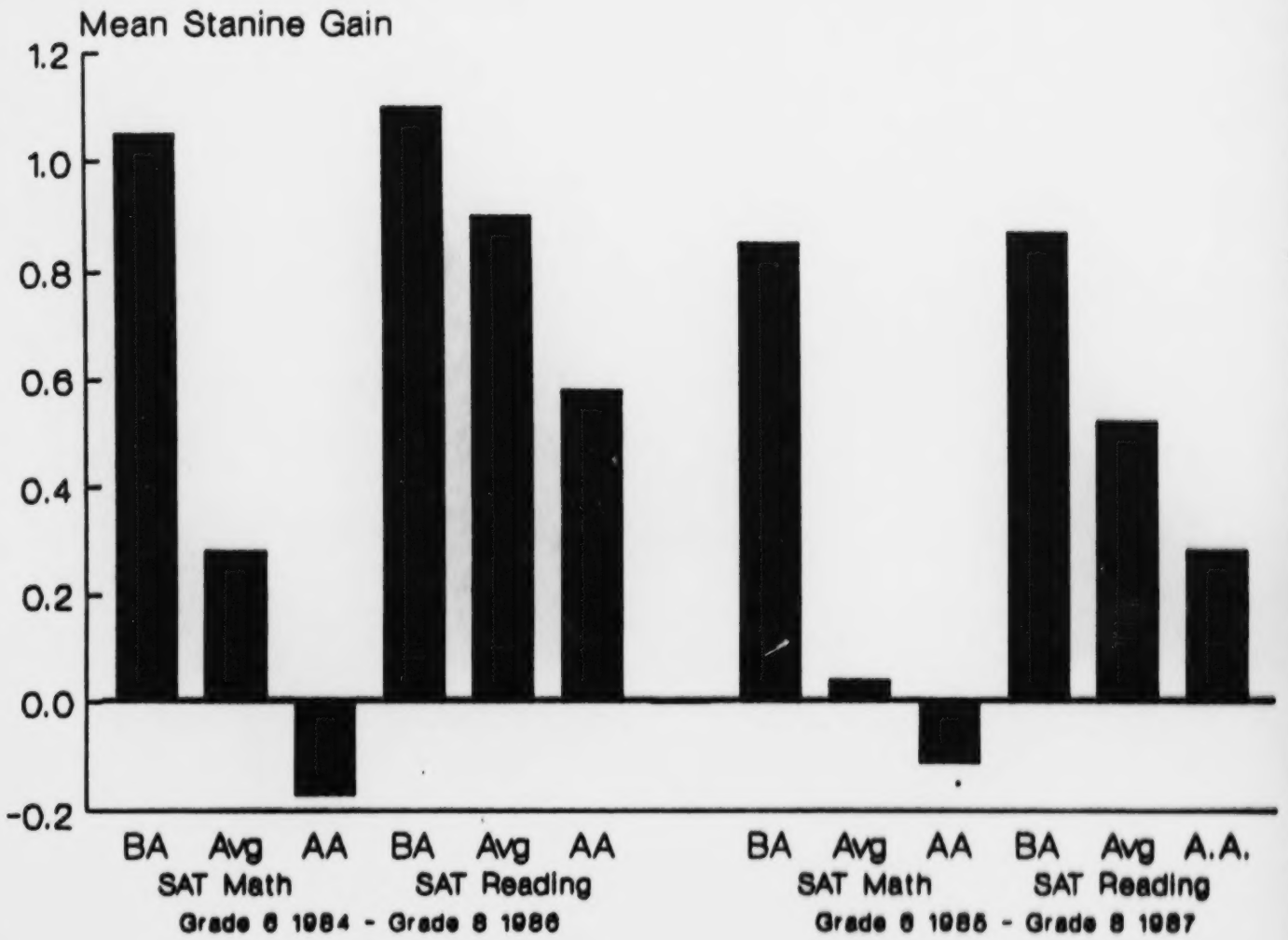


**Figure 6. Mean Gains by Pretest Achievment Level: Grade 8 to Grade 10**